J2
409/2:27
c.2

North Carolina State Library
Raleigh
N. C.
Doc.

SCHS STUDIES

A SPECIAL REPORT SERIES BY THE N.C. DEPARTMENT OF HUMAN RESOURCES, DIVISION OF HEALTH SERVICES, STATE CENTER FOR HEALTH STATISTICS, P.O. BOX 2091, RALEIGH, N.C. 27602

NOV 30 1983

No. 27 November, 1983

The Utilization of North Carolina General Hospitals: A Geographic Analysis

by

Michael J. Patetta

and

Paul A. Buescher

Abstract

The variation in general hospital utilization rates among the counties in North Carolina is depicted and analyzed. The principal statistical methods used are multiple regression and correlation analysis. It was found that the number of short-term general hospital beds per 1,000 population is the best predictor of county hospital utilization rates. The higher the beds per population in a county, the higher is the utilization of hospitals by residents of the county. Also, high hospital utilization is associated with a low number of physicians per 1,000 population. It is hoped that the results of the analysis will be useful to health planners and to those involved in the delivery of health care in a county. This article should also serve as a brief review of the use of multiple regression and correlation analysis in health care studies.

Contents

	Page
Introduction	2
Data Collection	2
Methodology	2
Results	3
Correlation Analysis	3
Factor Analysis	5
Regression Analysis	5
Residual Analysis	7
Summary and Discussion	7
References	10
Maps	11
Appendix	15

Introduction

In 1981, the highest North Carolina county hospital discharge rate was more than triple the lowest rate. The question arises as to what factors contribute to this wide variation. For example, why do Orange County residents have 73 discharges per 1,000 population while Avery County residents have 243 discharges per 1,000 population? Is it a matter of demographic structure of the population, more available hospital beds, or is it simply that the residents of Orange County are in better health than the residents of Avery County? These are the types of questions that this paper addresses.

The purpose of this paper is, therefore, to exhibit and explain the geographic pattern of general hospital utilization rates by county in North Carolina. These utilization rates are residence-based and represent how much the residents of a county utilize general hospitals in any location. Facility-based rates, which represent how much the hospitals in a county are being utilized, are not employed in this study and should not be confused with the residence-based rates.

Basically, the variables that affect hospital utilization can be divided into resident characteristics variables and medical resource variables. Examples of resident characteristics variables that have been found to impact hospital utilization are: health status of the population (11,13,20); patients' perceptions of the health care facility (18); educational attainment of the population (5); age of the population (4,11); distance from a health care facility (17); regular source of medical care (15,16); insurance coverage of the population (6,15,22); per capita income of the population (11,16); racial composition of the population (2,6); and sex composition of the population (9). Unfortunately, some of these variables were not used in the present study because they could not be easily quantified by the researchers.

Medical resource variables measure the supply of health care in an area. Examples of resource variables that have been found to influence hospital utilization are: bed to population ratio (1,6,22); surgeon to population ratio (6,23); inpatient occupancy rate (7); clinical decisions of physicians (12,23); and specialties of physicians in an area (23). Many of these variables were available and were employed in the present analysis.

Data Collection

The dependent variable, or the variable that needs to be explained, was formed using patient origin data. These data were obtained from the application for renewal of license to operate a hospital, administered by the N.C. Division of Facility Services. North Carolina non-federal hospitals are required to put on this application the number of patients admitted or discharged by county of residence and by age. A rate for residents of each county was derived by summing across hospitals the 1981 non-federal general hospital patients by county and then dividing each sum by the total county civilian population. These rates were broken down by age so that each county had four different rates: total discharges, ages 0-13, ages 14-64, and ages 65 and over. Finally, the rates for counties with a significant number of patients leaving North Carolina for hospital care were adjusted to estimate total utilization based on information from two previous special studies (19,21). It is assumed that the general patterns of outmigration observed in 1974 and 1978 were in operation in 1981. Maps of the discharge rates by county are shown in Figures 1 through 4.

The independent variables, or the variables used to explain variation in the dependent variables, were obtained from a variety of sources. In large part the resident-based variables were derived from the 1980 U.S. Census or from North Carolina vital statistics data. The facility-based or medical resource variables were derived from data sources such as the *North Carolina Health Manpower and Facilities Data Book* and the *Health Facilities Data Book* hospital volumes. A complete list of all the sources and the variables used in the analyses is shown in the Appendix.

Methodology

Once the variables were selected, a number of statistical tests were executed in order to examine the relationships among the variables. The Statistical Analysis System (SAS) package computer program was used for all of the analyses. The first statistical test used was correlation analysis, which measures the strength of the relationship between two variables. The Pearson product-moment correlation generates coefficients that range from –1 to +1. A correlation coefficient close to +1 means that the two variables are highly positively correlated (they are directly related); a correlation coefficient near zero means there is little correlation; and a correlation coefficient close to –1 means that the variables are highly negatively correlated (inversely related). Furthermore, squaring the correlation coefficient shows the percentage of the variation explained in one variable by the other variable.

Another technique used in the study was multiple regression analysis, which investigates the effects of several independent variables on a dependent variable (hospital utilization). It differs from correlation analysis in that it allows us to assess the unique contribution of each variable. One of the calculations produced in the analysis is the R^2 statistic, which indicates the proportion of the total variance in the dependent variable that is accounted for by all of the independent variables. An adjusted R^2 is an R^2 that has been adjusted for "degrees of freedom," that is, the number of cases in relation to the number of variables. This is important because when the number of variables used begins to approach the number of cases, the R^2 may be due to chance fluctuations. Therefore, the adjusted R^2 takes into account the number of variables and it should always be lower than the unadjusted R^2.

A major problem with multiple regression analysis is multi-collinearity, which occurs when the independent variables are highly correlated with one another. An example of this phenomenon is when education and income, which are both strong indices of socioeconomic status, are entered into the same regression equation. This could produce a bias in the equation since the assumed "independent" variables may measure very much the same thing. In order to help solve this problem, a statistical technique called factor analysis was employed.

Factor analysis basically groups variables into clusters or "factors" on the basis of their intercorrelations. The derived factors will be uncorrelated if a solution is selected that results in an angle of 90° between each factor. Factor analysis also generates a factor score on each factor for each observation. In this analysis, an observation was a county and a factor score indicated the strength of a county's measurement on a given factor. These factor scores were then used as independent variables in a multiple regression equation in order to see how much of the variation in the dependent variable (hospital utilization) is explained using the factors.

Another technique employed in the analysis was stepwise multiple regression. This method was used to select for further analysis the ten best medical resource and residence-based variables. Stepwise regression enters the independent variables into a regression equation by their order of importance in accounting for variance in the dependent variable. This ordering is determined by a type of "partial" correlation coefficient, which indicates the strength of the relationship between an independent variable and the dependent variable once the effects of the other independent variables have been accounted for. Stepwise regression does not, however, control for multi-collinearity (in the SAS program); therefore, only the independent variables with intercorrelations of less than .80 were entered into the procedure. When two of the independent variables were correlated at .80 or greater, indicating that they measure much the same thing, the one least correlated with the dependent variable was eliminated.

Results

Correlation Analysis

Results of the correlation analysis are shown in Table 1. One can see that the percent of the population that are Medicare disabled enrollees had the strongest relationship with three of the utilization rates. In the case of total discharges per 1,000 population, squaring the correlation coefficient shows that the percent of Medicare disabled enrollees explains 32 percent of the variation in hospital utilization. This strong association is somewhat curious when one considers that the Medicare disabled enrollee population is quite small (the percent of the total population that are Medicare disabled enrollees ranges from 0.7 to 2.5 across the counties). A possible explanation is that the percent of Medicare disabled enrollees is a surrogate measure for other variables that do have a direct impact on hospital utilization. For example, the percentage of Medicare disabled enrollees could be strongly correlated with the percentage of the population in poor health. Thus it could be a generally poor health status of the population causing high hospital utilization rather than the impact of just the group of Medicare disabled enrollees.

Other variables highly correlated with hospital utilization include average length of hospital stay, percent of population female age 15-44, hospital beds per 1,000 population, and percent of employees in the secondary sector (manufacturing, construction, etc.). The negative coefficient for length of stay suggests that when hospital patients stay in a hospital for a long period of time, the turnover rate, and therefore the discharge rate, is lower. The correlation for the employees in the secondary sector was an unexpected result since that has not often been documented in the literature. This result could be related to the composition of the North Carolina manufacturing sector, which is heavily concentrated in textiles and furniture. The correlation for beds to population ratio was not surprising, though, since this has been frequently documented in the literature. What was surprising was that it only had a high correlation with one

Table 1
The Top Five Correlations Between
Hospital Utilization Rates and Other Measured Variables

Utilization Rates	Related Variables	Correlation Coefficient	Significance*
Total Discharges per 1,000 Population	Percent of population Medicare disabled enrollees	.570	.0001
	Percent of population with work disability	.438	.0001
	Percent of population female age 15-44	-.427	.0001
	Average length of stay	-.417	.0001
	Percent of population Medicare aged enrollees	.386	.0001
Discharges age 0-13 per 1,000 population	Percent of population Medicare disabled enrollees	.250	.012
	Percent of employees in secondary sector	.244	.015
	Average length of hospital stay	-.242	.015
	Percent of physicians age 60 and over	.190	.057
	Percent of employees in tertiary sector	-.186	.063
Discharges age 14-64 per 1,000 population	Percent of population Medicare disabled enrollees	.557	.0001
	Average length of hospital stay	-.411	.0001
	Percent of population with work disability	.398	.0001
	Percent of employees in secondary sector	.397	.0001
	Percent of population female age 15-44	-.373	.0001
Discharges age 65 and over per 1,000 population	Average length of stay	-.432	.0001
	Beds per 1,000 population	.328	.0009
	Percent of patients leaving county for care	-.288	.0037
	Percent Medicare of total hospital discharges	.285	.0040
	Percent of population Medicare disabled enrollees	.256	.0102

*The significance level is basically the probability that the observed relationship could have occurred by chance.

of the utilization rates (discharges age 65 and over). The negative correlation of hospital utilization with percent females age 15-44 is probably because an area with a high percentage of females age 15-44 will have a relatively young population, which has low utilization. This will counteract the higher number of obstetric discharges from this group, since obstetric discharges are a small percent of the total.

Factor Analysis

The factor analysis, which was intended to group the independent variables into a smaller set of uncorrelated "factors" to use in regression analysis, produced five major interpretable factors: medical resource factor, poverty factor, elderly factor, rural factor, and a white-collar factor. The prediction of hospital utilization using these factors was, however, rather poor with only 11 percent of the variation in total hospital utilization rates accounted for. A possible explanation for this is that the factor analysis grouped independent variables based on their correlations among themselves, not on their correlations with the dependent variables. In general, however, factor analysis should be considered as a technique for reducing a set of intercorrelated variables prior to regression analysis.

Regression Analysis

Table 2 exhibits the results of multiple regression analysis. The ten best variables were first chosen by stepwise regression and then entered into a multiple regression procedure that generated standardized weights for the independent variables. (In the SAS computer program that was used, the stepwise procedure does not produce standardized weights.) These weights indicate how much change in the dependent variable is produced by a change in one of the independent variables when the others are statistically held constant. Since these weights are standardized, we can rank the variables by their ability to predict utilization, after the variance shared with all of the other independent variables in the equation has been removed. The results from Table 2, therefore, indicate that the bed-to-population ratio is the best single predictor of utilization in this analysis, for total utilization and two of the age groups. The higher the beds per population in a county, the higher is the utilization of hospitals by residents of the county. Another very important variable is the physician-to-population ratio, which has a high, **negative** association with three of the four utilization rates, meaning that a low number of physicians per 1,000 population is related to a high rate of hospital utilization.

Other findings in Table 2 include the consistently positive effect on utilization of the percent of the population that are Medicare disabled enrollees. As was stated before, the percent Medicare disabled enrollees could be highly correlated with other variables that have a more direct impact on utilization. Average length of stay, meanwhile, has a consistently inverse relationship to the utilization rates. This negative relationship was also found in the correlation analysis.

The negative weight for the percent Medicare of total resident patients is rather curious since the simple correlation between this variable and utilization is positive. Thus, even though an area with high hospital utilization is likely to have a high percentage of Medicare patients, the contribution of this variable to the prediction of utilization in the multiple regression equation is negative. This paradox results from the fact that in multiple regression the sign of the weight applies to the relationship between that portion of hospital utilization unexplained by the other nine variables and that portion of the percent Medicare of total resident patients that is unrelated to these other nine variables. In this case, the Medicare percentage is positively related to some of the other variables in the equation and the positive explanatory effect of the Medicare percentage has apparently already been "used up" by these other variables. The weights of some other variables in Table 2 may have signs that run counter to "common sense." But again, keep in mind that multiple regression maximizes the power of a **combination** of variables to predict utilization, and this may produce results different from what a series of two-variable relationships would indicate.

As was noted in the methodology section, the R^2 statistic indicates the proportion of the total variance in the dependent variable that is accounted for by all of the independent variables. Thus we can account for around 40 percent of the variation in discharges age 0-13 per 1,000 population and 55 percent of the variance in the age 65 and over rates. For ages 14-64 and the total discharge rate, over 60 percent of the variance is accounted for by the regression equations, which is a reasonably good level of prediction.

The F value in Table 2 measures the significance of R^2. It is a test to see if the R^2 may be due simply to random variation. The probability statistic (p) represents the probability of getting the observed F value simply by chance. In all four multiple regression equations, the probability statistic is 0.0001 or lower which means that

Table 2
Multiple Regression Analysis Summary:
Ten Best Predictors of Age-Specific and Total Hospital Utilization

1. Total Discharges per 1,000 Population

Independent Variable	Standardized Weight	
Beds per 1,000 population	.793	
Physicians per 1,000 population	-.342	
Average length of stay	-.336	R^2 = .665
Percent of population Medicare disabled enrollees	.335	Adjusted R^2 = .628
Percent of population, high school graduates	-.240	F = 17.693
Percent Medicaid of total resident patients	-.236	p = .0001
Hospital occupancy rate	.227	
Hospital outpatient visits per 1,000 population	-.220	
Health Department expenditures per 1,000 population	.173	
Percent of physicians age 60 and over	.130	

2. Discharges Age 0-13 per 1,000 Population

Independent Variable	Standardized Weight	
Beds per 1,000 population	.508	
Percent secondary sector employees	.478	
Percent Medicare of total resident patients	-.415	R^2 = .428
Percent primary sector employees	.381	Adjusted R^2 = .363
Hospital occupancy rate	.304	F = 6.650
Average length of stay	-.280	p = .0001
Percent of population Medicare disabled enrollees	.237	
Percent of physicians over age 60	.211	
Percent of population age 15-64	.189	
Infant mortality rate	-.157	

3. Discharges Age 14-64 per 1,000 Population

Independent Variable	Standardized Weight	
Beds per 1,000 population	.744	
Physicians per 1,000 population	-.398	
Percent of population Medicare disabled enrollees	.381	R^2 = .661
Percent Medicare of total resident patients	-.338	Adjusted R^2 = .623
Average length of stay	-.333	F = 17.360
Percent of population female age 15-44	-.310	p = .0001
Hospital occupancy rate	.192	
Percent of population, white	.180	
Health Department expenditures per 1,000 population	.156	
Hospital outpatient visits per 1,000 population	-.145	

4. Discharges Age 65 and Over per 1,000 Population

Independent Variable	Standardized Weight	
Average length of stay	-.536	
Beds per 1,000 population	.517	
Percent Medicare of total hospital discharges	.517	R^2 = .570
Physicians per 1,000 population	-.317	Adjusted R^2 = .522
Percent primary sector employees	.307	F = 11.811
Total number of hospitals in county	.270	p = .0001
Percent Medicare of total resident patients	-.243	
Nursing home beds per 1,000 population	.216	
Hospital occupancy rate	.214	
Percent of surgical visits that are inpatient	.152	

the probability of the observed associations between the dependent variable and the set of independent variables occurring by chance is extremely small.

Residual Analysis

While the multiple regression equations described above accounted for as much as 60 percent of the variation in county-level utilization, this still leaves at least 40 percent unexplained. Thus one might ask for which counties is the prediction of utilization good and for which counties is it poor. A procedure that helps answer this question is the mapping of regression residuals. A residual is basically the difference between the actual value and the predicted value generated by the regression equation. In this analysis, the residuals identify the counties that have utilization rates that are not being predicted very well by the multiple regression equation. This could lead to the selection of areas for intensive investigation and to the formation of additional variables to use in explaining utilization.

The regression residual maps are shown in Figures 5 through 8. One county that stands out in three of the maps is Gates County. The residual for this county is more than two standard errors above the predicted value, which indicates that the residents are utilizing hospitals much more than would be predicted based on its values on the independent variables. A possible reason for this discrepancy is that some counties along the border had their discharge rates adjusted for an estimated percentage going out of state for hospital care. If this adjustment were too high for Gates County, then that county's discharge rate would be artifically inflated. Another possible reason for Gates County having a high residual is that the county's population is very small. Small populations produce a small number of discharges in the numerator, and artifically high or low rates may occur in a random fashion. However, if a county has a large population and little migration out of North Carolina for hospital care, more confidence can be placed in the measured utilization rates, and thus the residuals will be better indicators of unusual county utilization.

Another technique that may suggest additional predictor variables to include is spatial autocorrelation. This procedure tests whether values in one geographic place are dependent on values in another geographic place. Positive spatial autocorrelation exists when the values in adjacent places are similar. This signifies a geographic clustering of similar values. There is no autocorrelation when the values of the places are randomly distributed.

In this analysis, residuals from the regression equation were entered into the spatial autocorrelation test. Spatial autocorrelation of the residuals would suggest the existence of spatial or regional explanations for the pattern of deviations from the predicted values. The residuals were first divided into positive and negative values, and the number of adjacencies between the positive residual counties and the negative residual counties were tested against a random distribution. Chi-square statistics were then calculated to evaluate whether or not the frequencies that were empirically obtained differed significantly from the expected values. For all four maps the chi-square statistic was not significant. Consequently, the residuals from the regression equation do not appear to be spatially autocorrelated.

Summary and Discussion

Several statistical analyses were used to assess the relative impact of various factors on general hospital utilization. The results showed that sixty-seven percent of the variation in total hospital utilization was explained by ten health resource and residence-based variables. Maps of counties with utilization rates substantially higher or lower than the rate predicted from the regression equation are presented in order to identify areas with unusual patterns and suggest reasons for these variations. For all but one age category, the bed-to-population ratio explained the largest proportion of variation in hospital utilization: in general, the higher this ratio in a county, the higher the utilization of its residents. Other important variables were average length of stay (negative relationship), the percentage of the population that were Medicare disabled enrollees (positive), and the physician-to-population ratio (negative).

The strong relationship between the bed-to-population ratio and inpatient hospital utilization has been documented in other studies (1,6,22). Even after holding constant the effects of other variables upon utilization, high bed supply has a positive impact. One explanation could be that empty hospital beds influence physicians' decisions to admit patients into the hospital. This would be a case of the supply of medical resources in an area influencing demand. Davis and Russell (7) found that the demand for outpatient care is sensitive to the inpatient occupancy rate. They suggest that when hospital inpatient facilities are crowded, physicians switch more patients into outpatient care; therefore, a policy aimed at restricting the supply of beds and keeping occupancy rates high may cause a reallocation of resources to less

costly forms of treatment. There is a question, however, as to which came first, a high bed-to-population ratio or the high utilization. If the high bed-to-population ratio preceded the high hospital utilization rate, then the explanation of supply influencing demand would receive support. If however, previously high utilization led to an increase in the bed ratio, then the need or demand for hospital services would seem to be influencing the bed supply. Either explanation would be consistent with the relationship observed here, and the direction of the influence may be different across counties.

The tremendous variation in utilization rates across the 100 counties in North Carolina is not likely to be due only to differences in need. If the second explanation above (high utilization rates come first and generate pressures for more beds) is appropriate, then differences in physician practice patterns probably account for some of the observed variation in the bed-to-population ratio. The relationship between the practice style of physicians and hospital utilization has been documented in other studies (23). For example, it has been suggested that physicians in some areas tend to hospitalize for conditions handled elsewhere more frequently on an outpatient basis. Unfortunately, there was no convenient way to measure this variable for the present study.

It should be mentioned that the data used here do not necessarily indicate that high utilization in an area with a high bed supply is inappropriate, and it is likely that in some areas the bed supply is less than what is needed. A possible alternative explanation for the positive relationship between bed supply and utilization is that North Carolina has fewer hospital beds than needed, and that more beds in areas of low supply would raise utilization to an appropriate level.

To further explore the relationship between the bed supply and hospital utilization, the analysis was repeated using total patient days per 1,000 population as the dependent variable (rather than discharges). The strongest single predictor was again the bed-to-population ratio, with a standardized weight nearly twice as large as the next most important variable. The bed ratio was also the strongest predictor of age-standardized utilization rates, where the age-specific rates of each county were applied to the state's population age distribution.

In a recent study by Getts (9) of hospital use at the Health Service Area (HSA) level, the bed-to-population ratio came out as the top predictor of "deviant" patient days per 1,000 population, i.e., patient days over or under those **expected** if U.S. rates are applied to the HSA population structure. When this approach was repeated using North Carolina data, the bed ratio was also one of the top predictors. Furthermore, the number of physicians per 1,000 population had a high negative coefficient in the present study and in Getts' study. All of the variables in our regression equation accounted for 58 percent of the variation among N.C. counties in deviant patient days, while in the Getts study among HSA's the percent was 79.

Part of the positive relationship of the bed ratio to utilization is due to 17 rural counties with no hospital beds, many of which also have very low utilization. In some of these areas hospital use may not be adequate for the needs of the population, and thus it is not just overuse in heavily bedded areas that contributes to this positive relationship. However, when the regression for total discharges per 1,000 population was repeated leaving out these 17 counties with no hospital beds, the bed-to-population ratio was still by far the most important predictor, with a strong positive relationship to utilization.

It should be noted that in all regressions using the **discharge** rate, average length of stay had a negative impact, indicating that long hospital stays result in a lower turnover rate, or perhaps that a low admission rate results in longer lengths of stay. With **patient days** per 1,000 population as the dependent variable, however, length of stay had very little association, which suggests that factors besides a long average length of stay contribute to a high patient-day use rate.

The negative relationship of physicians per 1,000 population to inpatient hospital utilization is a very interesting finding. This relationship remains even after adjusting the values of two extreme outliers.[1] Contrary to the idea that more physicians means more hospital admissions, it appears that many physicians are providing services that keep people out of the hospital. A lack of physicians in a county may cause residents to go to a hospital for many of their ailments. Another contributing factor may be that more physicians means more peer review which would affect physician admitting practices in a county and result in a higher proportion of certain conditions being treated on an outpatient basis.

---

[1]*Durham and Orange counties each have a medical school and thus a very high physician to population ratio, but both have relatively low resident utilization. These two unusual cases were strongly affecting the relationship, as revealed by a plot of the two variables, and thus for these two counties the physician-to-population ratio was recoded to the state average before the regression analysis was done.*

This relationship has important implications and deserves further study.

It should be noted that while regression analysis is a useful method for ranking the importance of a series of variables as they impact a "dependent" variable, there are certain potential limitations that must be kept in mind. While regression can show that an association exists, it is not safe to conclude causality just on this basis. Though the results may be suggestive, regression analysis **alone** does not demonstrate that a change in one independent variable directly leads to change in the dependent one. Regression may demonstrate coexistence, but causality must be imputed on theoretrical grounds or demonstrated by experimental methods.

Linear multiple regression, the type employed here, also involves certain assumptions; and if serious departures from these assumptions occur, the results may be biased. First, each of the dependent variables was plotted and found to approximate a normal distribution. We tested the assumption of a linear relationship between each independent variable and total discharges per 1,000 population by plotting the pairs of variables and found no serious deviations from linearity. Other assumptions are that the residual or error term (actual value minus predicted value) is uncorrelated with other variables in the equation and that the error variance is constant across different values of the other variables. Plots of the residuals against the predicted values, the sequence of cases, and each independent variable revealed that these assumptions were not violated.

In an earlier study in this series (5), a suggested approach was to predict hospital use rates based on demographic and need factors and then to determine what other variables account for deviations from these expected "need" values. As the present study developed, the difficulty of isolating one set of variables and saying a priori that **these** indicate need became more apparent. This is particularly true given the sometimes high correlations between the residence-based variables and the medical resource variables. Thus no attempt was made here to isolate a set of need indicators, but rather all major variables were included in the regression model to compare their importance, and deviations from the resulting predicted values were examined. The approach of quantitatively predicting a "needed" level of hospital use should, however, be pursued in future research.

While our analysis accounted for around 60 percent of the variation in hospital utilization, 40 percent remains unexplained by the variables that we could measure in a county-level analysis. One important variable not in this study is health insurance coverage. It has been well documented that lack of health insurance coverage is associated with low hospital utilization rates (6,15,22); therefore, a measure of the percent of county residents without hospital insurance could add to our ability to predict hospital utilization. Another variable that impacts hospital utilization is accessibility of hospital care (14). Dimensions of accessibility would include availability, travel distance, ability to pay, and acceptability of the services offered. Variations from county to county in the health status of the population should also account for some of this unexplained variation, though this dimension would be very expensive to quantify.

In conclusion, it is hoped that the results of this study will be useful to health planners and those involved in the delivery of health care in a county. Examining a county's values on variables found important in the prediction of hospital use may help explain unusually high or low levels of use. Furthermore, for counties with a much higher or lower level of use than would be expected given its values on the predictor variables (i.e., for counties with large residuals), further investigation may uncover reasons for these unusual patterns of hospital utilization that we were not able to measure. It should, however, be noted that even though a county's hospital utilization rate deviates from an average or expected rate, it does not necessarily mean that something is wrong. A low utilization rate could be due to less than optimal utilization, but it could also be due to very little unnecessary utilization or low morbidity. On the other hand, a high utilization rate could result from a high level of morbidity in that county. Questions about unusually high or low utilization can best be answered by those familiar with the situation in a particular county.

References

1. Anderson, J.G., "Demographic Factors Affecting Health Services Utilization: A Causal Model," *Medical Care.* Vol. 11, 1973, pp: 104-120.

2. Bashshar, Shannon, G.W. Metzner, C.A., "Some Ecological Differentials in the Use of Medical Services," *Health Services Research.* Vol. 6, No. 1, 1971, pp. 61-75.

3. Berkanovic, E., Telesky, C., and Reeder, S. "Structural and Social Psychological Factors in the Decision to Seek Medical Care for Symptoms," *Medical Care.* Vol. 19, No. 7, 1981, pp. 693-709.

4. Borhani, N.O. and Kraus, J.F., "Use of Health Services in a Rural Community," *Health Services Reports.* Vol. 88, No. 3, 1973 pp. 275-288.

5. Buescher, Paul, "1980 Hospital Inpatient Utilization by County for Medicare and Medicaid Patients," *SCHS Studies.* No. 24, February 1983.

6. Chiswick, Barry R., "Hospital Utilization: An Analysis of SMSA Differences in Occupancy Rates, Admission Rates, and Bed Rates," *Explorations in Economic Research.* Vol. 3, No. 3, Summer 1976, pp. 326-378.

7. Davis, K. and Russell, L.B., "The Substitution of Hospital Outpatient Care for Inpatient Care," *Review of Economics and Statistics.* Vol. 54, No. 2, 1972, pp. 109-120.

8. Dobson, A., Scharff, J., and Corder, L., "Six Months of Medicaid Data: A Summary from the National Medical Care Utilization and Expenditure Survey," *Health Care Financing Review.* Vol. 4, No. 3, March 1983, pp. 115-121.

9. Getts, C.V., "Hospital Use: Health Status and Health Systems Report Series," Institute for Health Planning, Madison, Wisconsin, June 1982.

10. Hassinger, E., and Hobbs, D., "The Relation of Community Context to Utilization of Health Services in a Rural Area," *Medical Care.* Vol. 11, No. 6, 1973, pp. 509-522.

11. Joseph, H., "Empirical Research on the Demand for Health Care," *Inquiry.* Vol. 8, No. 1, March 1971, pp. 61-71.

12. Lewis, C.E., "Variations in the Incidence of Surgery," *The New England Journal of Medicine.* Vol. 281, No. 16, October 16, 1969, pp. 880-885.

13. Luft, H.S., Hershey, J.C. and Morrell, J. "Factors Affecting the Use of Physician Services in a Rural Community," *American Journal of Public Health.* Vol. 66, No. 9, 1976, pp. 865-871.

14. Penchansky, R. and Thomas, J.W. "The Concept of Access: Definition and Relationship to Consumer Satisfaction," *Medical Care.* Vol. 19, No. 2, pp. 127-140.

15. Okada, L.M., "Impact of Community Health Centers and Medicaid on the Use of Health Services," *Public Health Reports,* Vol. 95, No. 6, 1980, pp. 520-534.

16. Rundall, T.G. and Wheeler, J.R., "The Effect of Income on Use of Preventive Care: An Evaluation of Alternative Explanations," *Journal of Health and Social Behavior.* Vol. 20, 1979, pp. 367-406.

17. Simon, J.L., "Change in Location of a Student Health Service," *Medical Care.* Vol. 11, No. I, 1973, pp. 59-67.

18. Snider, E.L., "Awareness and Use of Health Services by the Elderly," *Medical Care.* Vol. 18, No. 12, December 1980, pp. 1177-1182.

19. State Health Planning and Development Agency, Hospital Discharge Data, October 1978.

20. Tanner, James, "Predicting Physician Utilization," Medical Care. Vol. 21, No. 3, March 1983, pp. 360-364.

21. University of North Carolina, Area Health Centers Program, *Inpatient Utilization of General Hospitals in North Carolina,* October 1974.

22. Van der Gaag, J., Rutten, F. and Van Praag, B., "Determinants of Hospital Utilization in the Netherlands," *Health Services Research,* Vol. 10, No. 3, Fall 1975, pp. 264-277.

23. Wennberg, J. and Gittelsohn, A., "Variations in Medical Care Among Small Areas," *Scientific American.* Vol. 264, No. 4, April 1982, pp. 120-133.

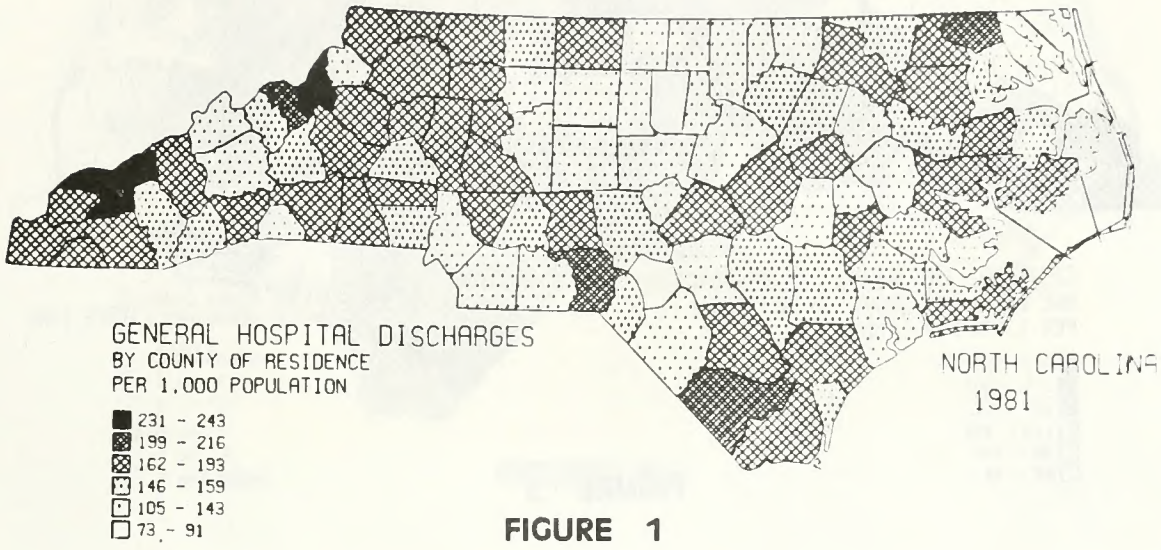

FIGURE 1

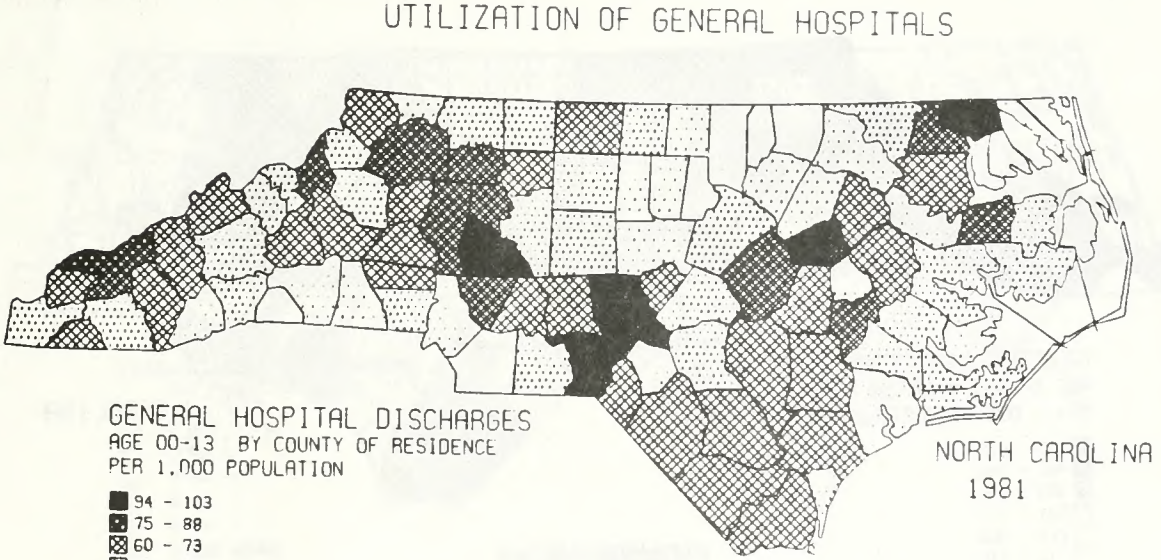

FIGURE 2

UTILIZATION OF GENERAL HOSPITALS

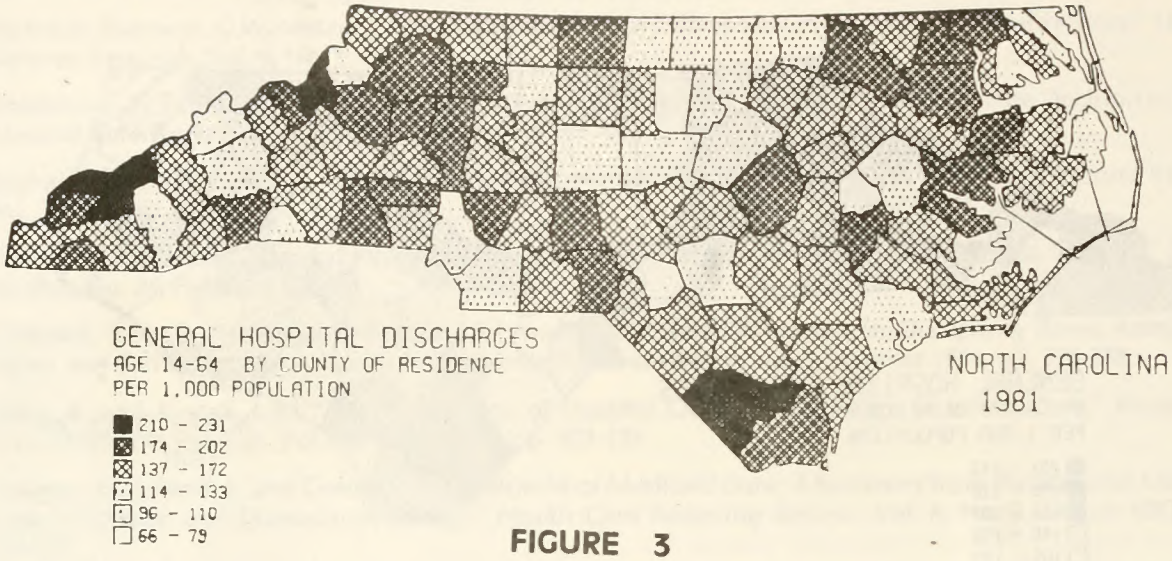

FIGURE 3

UTILIZATION OF GENERAL HOSPITALS

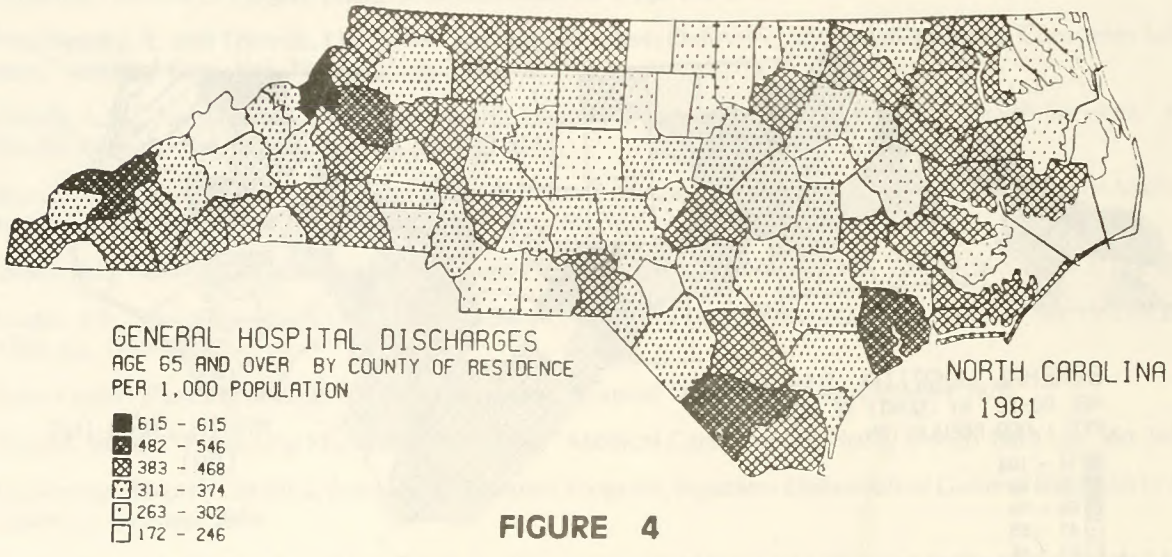

FIGURE 4

RESIDUALS FROM REGRESSION: TOTAL

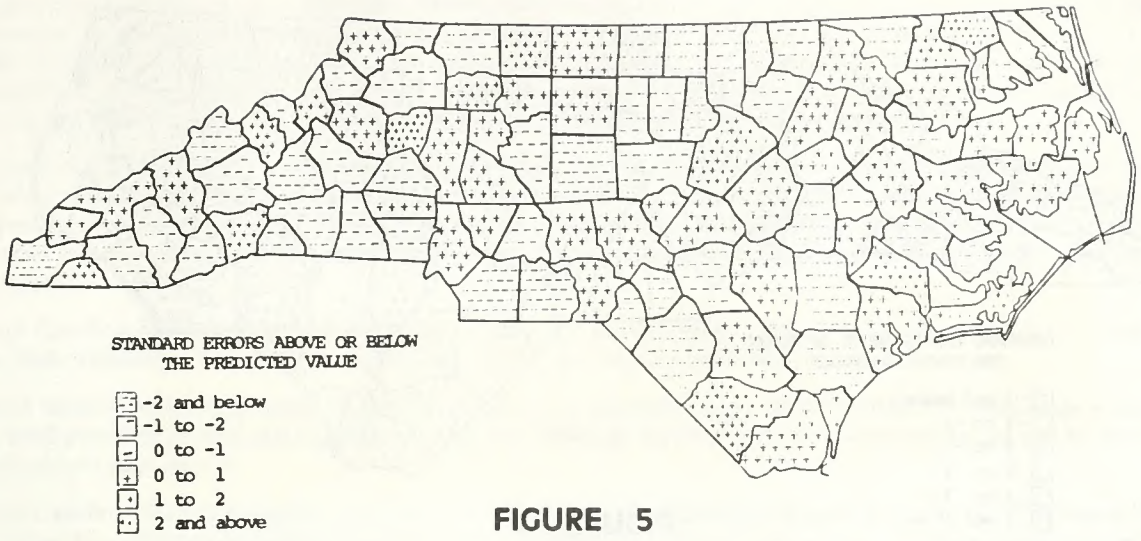

FIGURE 5

RESIDUALS FROM REGRESSION: 00-13

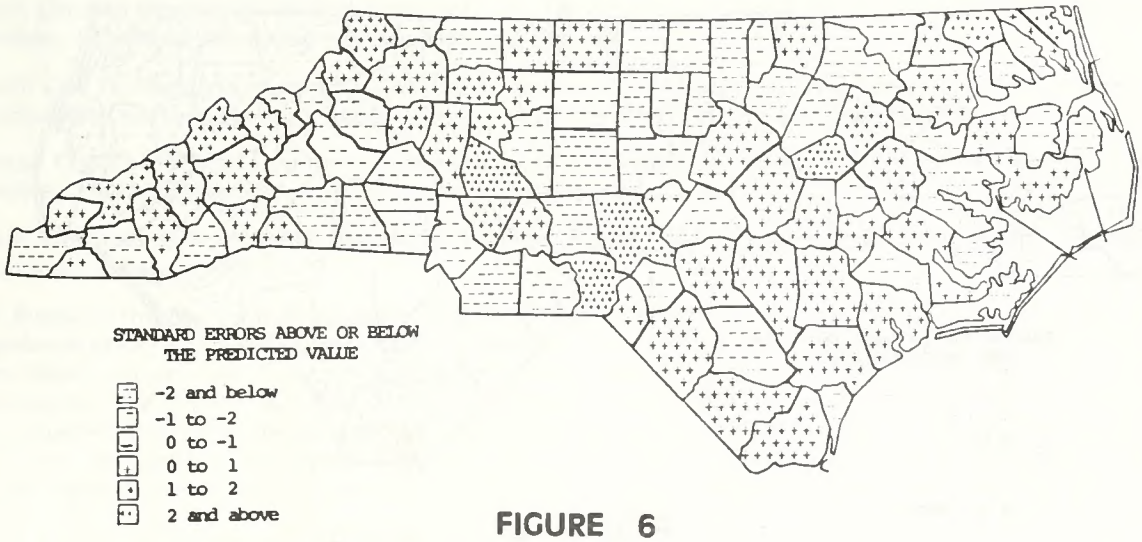

FIGURE 6

RESIDUALS FROM REGRESSION: 14-64

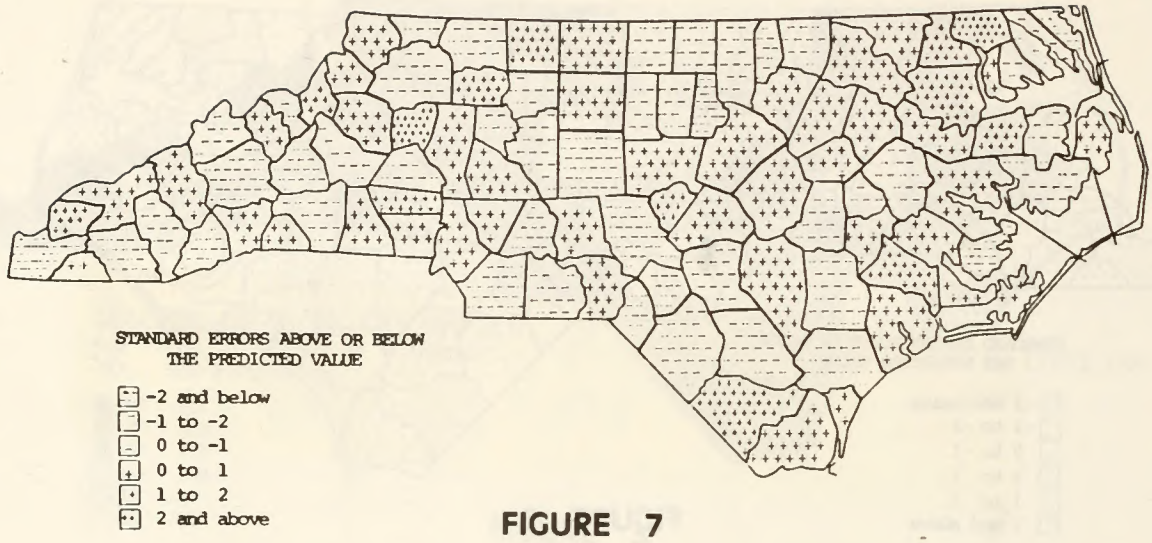

FIGURE 7

RESIDUALS FROM REGRESSION: 65 & OVER

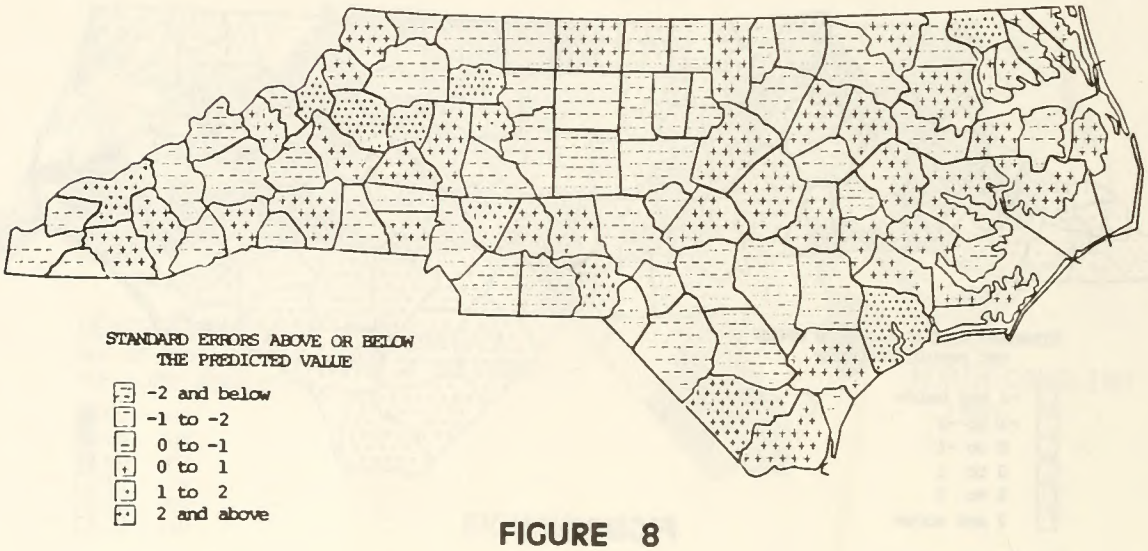

FIGURE 8

Appendix

List of Data Sources and Variables for Hospital Utilization Study

1. *Hospital Patient Origin Report, 1981 Data,* Health Facilities Data Book, N.C. State Center for Health Statistics (based on data from "Application for Renewal of License to Operate a Hospital, 1982", N.C. Division of Facility Services) October 1, 1980—September 30, 1981. Variables: Discharges age 0-13 per 1,000 population; Discharges age 14-64 per 1,000 population; Discharges age 65 and over per 1,000 population; Total Discharges per 1,000 population; Total patient days of care per 1,000 population; and Percent of patients leaving county for hospital care.

2. *Hospital Summary Report, 1981 Data,* Health Facilities Data Book, N.C. State Center for Health Statistics (based on Division of Facility Services' Licensure data), October 1, 1980—September 30, 1981. Variables: Non-federal short-term general hospital beds per 1,000 civilian population; Number of different specialties of physicians on hospital staffs; Total number of hospitals in county; Hospital outpatient visits per 1,000 population; and Percent of surgical visits that were inpatient.

3. *North Carolina Active Nonfederal Physicians,* Health Services Research Center, University of North Carolina, Chapel Hill, 1980. Variables: Physicians per 1,000 population and Surgeons per 1,000 population.

4. Health Services Research Center, Computer printout on physicians by age, 1980. Variables: Physicians age 40 and under per 1,000 population; Physicians age 60 and over per 1,000 population; Percent of physicians under age 40; and Percent of physicians over age 60.

5. *North Carolina Health Manpower and Facilities Data Book,* Health Services Research Center, UNC, Chapel Hill, 1980. Variables: Non-Primary care physicians per 1,000 population; Primary care physicians per 1,000 population; and Primary care physicians plus physician assistants and nurse practitioners per 1,000 population.

6. *County Health Data Book,* N.C. State Center for Health Statistics, 1981. Variables: Average occupancy rate and Length of stay for hospitals in county (from Division of Facility Services hospital licensure data); Infant mortality rate (1977-81).

7. Hospital Discharge Data, collected by North Carolina Hospital Association and the N.C. State Center for Health Statistics, 1980. Variables: Percent Medicare of total hospital discharges; Percent Medicaid of total hospital discharges; Percent Medicaid of total resident patients; and Percent Medicare of total resident patients.

8. *Nursing Home Summary Report, 1981 Data,* Health Facilities Data Book, N.C. State Center for Health Statistics (based on Division of Facility Services' Licensure data), November 1, 1980—October 31, 1981. Variables: Nursing home beds per 1,000 population.

9. North Carolina Division of Medical Assistance, N.C. Department of Human Resources, July 1, 1979—June 30, 1980. Variables: Percent of population who are Medicaid eligibles.

10. Health Care Financing Administration, U.S. Department of Health and Human Services, July 1, 1980. Variables: Percent of population who are Medicare aged enrollees; and Percent of population who are Medicare disabled enrollees.

11. Annual County Financial Information, reported to N.C. Division of Health Services, July 1, 1980—June 30, 1981. Variables: Health department expenditures per 1,000 population.

12. Public Revenues from Alcoholic Beverages, N.C. ABC Boards, July 1, 1980—June 30, 1981. Variables: Gross sales of alcohol per 1,000 population.

13. U.S. Bureau of the Census, 1980. Variables: Percent of population age 0-14; Percent of population age 15-64; Percent of population age 65 and over; Percent of population age 75 and over; Percent of population who are white; Percent of population who are male; Percent of population who are female age 15-44; Percent of employees in primary sector (farming, forestry, mining, etc.); Percent of employees in secondary sector (manufacturing, construction, etc.); Percent of employees in tertiary sector (administrative, service, etc.); Percent of population age 25 and over who are high school graduates; Per capita income (1979); Unemployment rate; Percent of population living in urban areas; Percent of population with a work disability; and Percent of population below the poverty level (1979).

14. *Leading Causes of Mortality, N.C. Vital Statistics, Volume 2, 1981,* N.C. State Center for Health Statistics, Variables: Crude death rate and Adjusted death rate (1979-81).

15. *North Carolina Vital Statistics, Volume 1, 1980,* N.C. State Center for Health Statistics, Variables: Age 0-14 mortality rate; Age 15-64 mortality rate; and Age 65 and over mortality rate.

16. *North Carolina People 1981,* N.C. Department of Human Resources, Title XX Section (All of these variables are synthetic estimates.) Variables: Percent of household heads in a low skill occupation; Percent of households in unsanitary living conditions; Percent of households with inadequate nutrition; Percent of households having difficulty obtaining medical attention; Percent of households that have one member with a serious medical problem; Percent of households that have the potential for health problems; Percent of households that have one member who has had difficulty obtaining transportation; and Percent of households living in substandard housing.

STATE LIBRARY OF NORTH CAROLINA
3 3091 00753 7145

Department of Human Resources
Division of Health Services
State Center for Health Statistics
P.O. Box 2091
Raleigh, N.C. 27602-2091
919/733-4728

BULK RATE
U.S. Postage PAID
Raleigh, N.C. 27602-2091
Permit No. 1862

CHERYL W. MCLEAN
ASST. DOCUMENTS LIBRARIAN
DIVISION OF STATE LIBRARY
ARCHIVES/LIBRARY BLDG.

N.C. Division of Health Services
State Center for Health Statistics
11-83/625/4¢/copy